

A Merchant's Guide to Identity Verification in INDIA

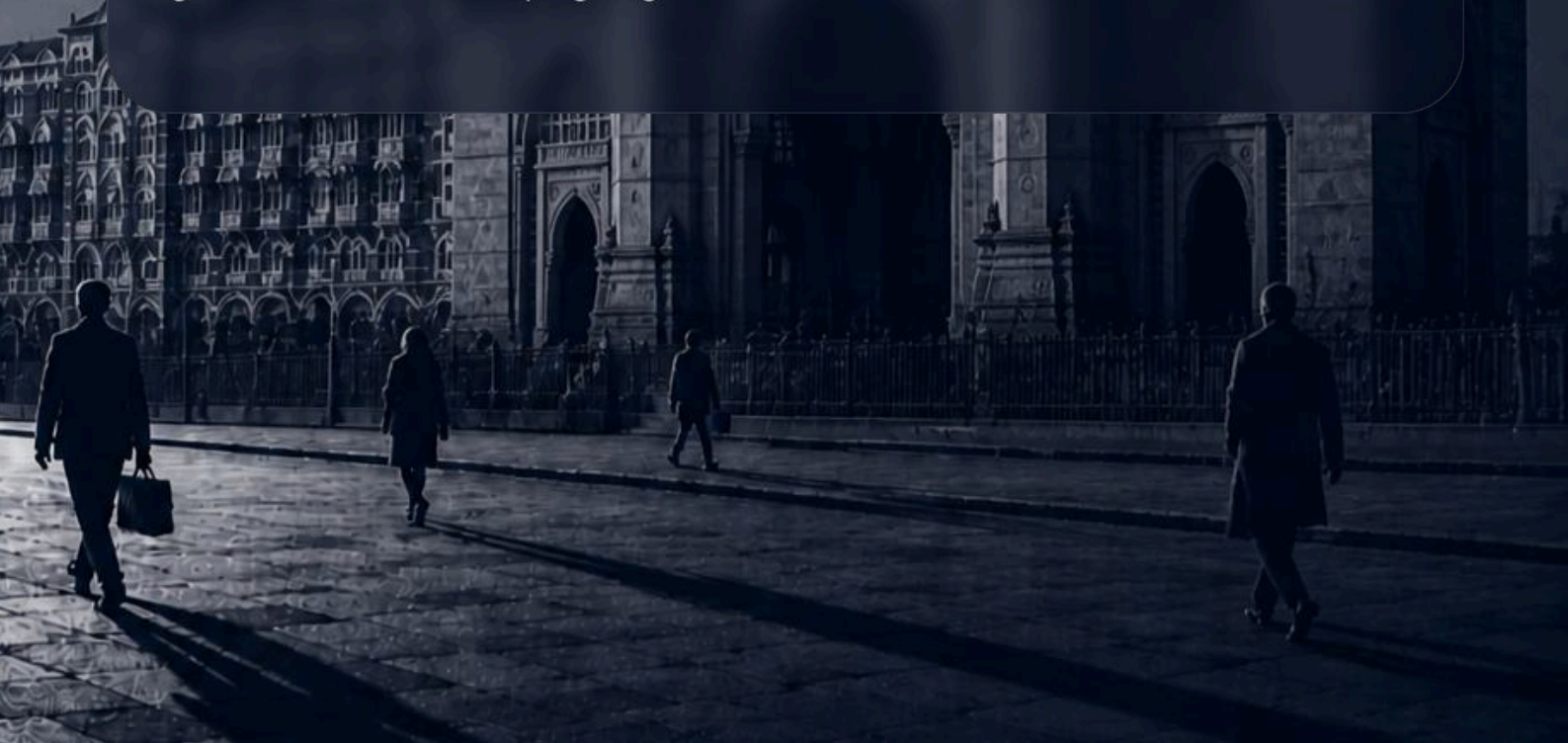
✓ Identity Verified

Real-Time Checks

Introduction

A customer in Mumbai can prove who they are with an Aadhaar letter printed on plain paper, an e-Aadhaar PDF pulled from a phone, or a laminated voter card issued twenty years ago. A customer in Chennai can hold the same proofs in Tamil rather than Hindi. All of them are genuine, and your verification flow has to read every one. India issues identity documents through central authorities, 28 states, and 8 union territories, and its constitution lists 22 scheduled languages, most with a script of their own.

This guide gives you the working knowledge to verify Indian customers with confidence. It covers the documents in circulation and the scripts that trip up automated reads, the data protection and KYC regime you verify under, which documents perform best for extraction, and the settings that lift pass rates for genuine users while keeping forged and borrowed documents out.



Understanding the Challenges of India's Identity Documents

Six traits of Indian documents that decide whether your automated checks read them or reject them.

Reliable verification in India starts with the documents themselves. They arrive in more than a dozen scripts, in formats that differ by issuing state, and in conditions that range from crisp PVC cards to laminated paper worn soft at the folds. The six challenges below cause most failed reads and false rejections.

1 | Linguistic Diversity

Indian identity documents carry text in many writing systems, including Devanagari, Bengali, Tamil, Gujarati, Telugu, Kannada, Malayalam, Gurmukhi, Odia, and Urdu. The Constitution of India lists 22 scheduled languages, and a single document often mixes two or three of them. An Aadhaar card prints in English alongside the holder's regional language, and voter cards follow the language of the issuing state. Shufti's OCR is trained across these scripts and reads documents in 150+ languages, so names and addresses extract correctly from an Aadhaar card, a voter ID card, or a driver's license whichever combination appears.



A voter ID card, front, printed in Tamil and English.



The back of the same card carries the address and instructions in Tamil.

2 | Complex Scripts with Vowel Modifications

Scripts such as Devanagari (नमस्ते), Tamil (நமஸ்தே), and Malayalam (നമസ്തേ) attach vowel marks to consonants and build compound characters, so one visual cluster can hold several sounds. A standard OCR engine reads these clusters as separate symbols and returns broken names. Shufti's AI-driven OCR detects and interprets the marks as part of the character they modify, which keeps extraction accurate on documents written in these scripts.

3 | No Spaces Between Words in Some Scripts

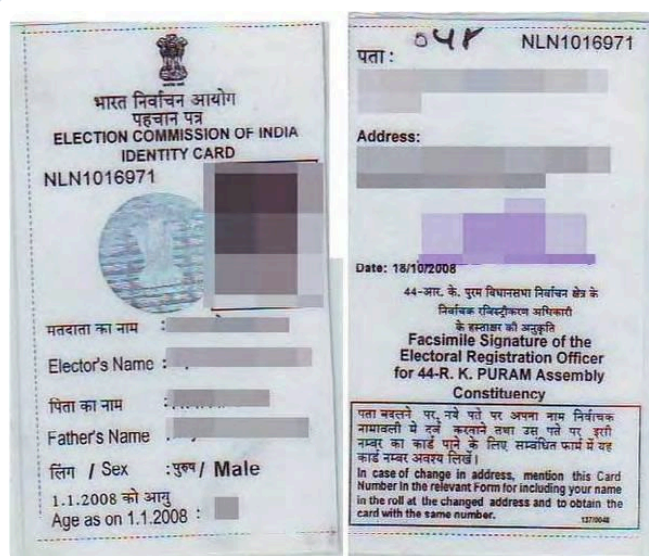
Written Tamil and Telugu often run words together without the spacing an OCR engine expects, so a name and an address can arrive as one unbroken string. Shufti's context-aware OCR segments these strings correctly and extracts each field despite the missing spaces.

Bangla	অনেকগুলি প্রাচীর মধ্যে পশ্চিমী মানসিকতার গতিপ্রবাহ
Devanagari	दोपहर का समय था। मि० काबई नाथता करके सिंगार
Gurmukhi	ਤਰ੍ਹਾਂ ਭੁੱਟਿਆ ਜਾਂਦਾ ਜਿਸ ਨਾਲ ਹੋਸ਼ ਵੱਧ ਵੱਧ ਹੋ ਜਾਂਦੇ
Oriya	ଚଢ଼େ ଗଢ଼ିତ୍ତା ତାଗ । ହନୁହାଡ଼ ଦୁଇଗା ଗଢ଼େଇ ହେଲାପରି
Gujarati	રહે છે, તેની સ્વધર્મ આશયી. પછી સમય એવો આવ્યો
Tamil	தோடு இரண்டறக் கலந்து விடவேண்டும்.
Malayalam	കൂട്ടരേ, നിങ്ങളൊക്കെയൊക്കെയും ഒരു വിവരമറിഞ്ഞുവോ?
Telugu	సిస్టమ్ నిలబడిపోయారు. పదండి! అందరం బైరానీని చూసి
Kannada	ಕೊರಗಿಸದೆ, ಸಲಹೆದ ಕಾಲುಮೆಯಿದು ಎಂದು ಪ್ರೇಮದಿಂದ ಆದರ
Urdu	حضرت امیر المومنین علیہ السلام کے وہ خطبات ارشادات اجتماعات اور کلمات قصار
English	For more than 30 years, scientists have been

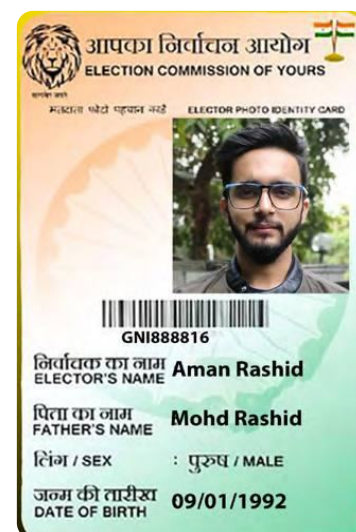
Eleven of the writing systems an OCR engine meets on Indian documents, from Bangla to Urdu.

4 | Cultural and Regional Variability

Identity documents vary across India's regions and languages. Voter cards are the clearest example, since their format depends on the issuing state and older laminated paper cards still circulate beside newer card versions. Shufti's continuously updated OCR recognises these regional variations and processes old and new formats side by side.



An older paper-format voter identity card, front and back, of the kind still presented for verification.



A newer elector photo identity card in the upright card format with a barcode and photograph.

5 | Driver's License Variations

Driver's licenses in India are issued by state transport departments, and design, language, and layout differ from state to state. A Maharashtra license, a Gujarat license, and a Delhi license share the same legal force but little of the same look, and most carry bilingual or trilingual text in English, Hindi, and the state language. Shufti's OCR is updated continuously against these variations, so names, addresses, and license numbers extract accurately from every state's format.



A Maharashtra State motor driving licence, smart card format with a chip.



A Gujarat State driving licence, a different layout in the same class.



A Delhi Transport Department licence, one more state format.

6 | Poor Document Quality

Documents submitted from rural areas often arrive as photocopies, faded originals, or poorly lit phone captures, and lamination adds glare on top. Text extraction suffers on all of these. Shufti's AI models are trained on low-resolution and degraded images, which lifts read accuracy on documents that would fail a standard check.

A Merchant's Guide to Identity Verification in India

The rules you verify under, the documents you can accept, and how each one performs. India's data protection regime is the Digital Personal Data Protection Act, 2023, operationalised through the Digital Personal Data Protection Rules, 2025 and overseen by the Data Protection Board of India. If you onboard customers in regulated sectors you also work under the Prevention of Money Laundering Act, 2002 and the Reserve Bank of India's KYC Master Direction, which sets out the officially valid documents a customer may present. Verification that mishandles identity data risks penalties under these laws and the harder cost of customers who stop trusting you with their documents.

Compliance and conversion are the same problem in India. A flow that asks for the wrong documents or stalls at capture loses customers mid-signup, so a compliant check that passes genuine users first time protects your regulatory standing and your completion rate together.

2.1 | Verification Modes

Shufti offers two verification modes; the right choice depends on how much of the capture journey you want to own.

ONSITE VERIFICATION

Runs through Shufti's iFrame and Verify component: your customer submits identity and address proofs directly and gets real-time guidance at each step. Results reach you in the Back Office portal or through the API the moment the check completes. Best for high-risk, compliance-heavy flows.

OFFSITE VERIFICATION

Puts you in charge of collection: you gather the proofs and submit them to Shufti, which suits capture inside your own app or at a counter. Results arrive the same way, in the Back Office portal or through the API. Fits telecom, e-commerce, and micro-finance batch checks.

2.2 | Document Types You Can Accept

Indian customers can prove identity with several documents, each offering a different level of security and extraction value. Specimens appear beside the performance grid in 2.3.

ID Card. India's national-ID family — four types circulate, each at a different level of security.

Aadhaar. The most widely used ID in India, issued by UIDAI with a unique 12-digit number linked to biometric data. Circulates as a paper letter, an e-Aadhaar PDF, and a PVC card.

PAN Card. From the Income Tax Department, primarily for tax but a reliable identity proof in financial services. Carries a 10-character alphanumeric code and a photograph.

Voter ID Card. Issued by the Election Commission of India with a photograph and alphanumeric number. No biometric security, so treat it as a supporting document, not a primary one.

Ration Card. Issued to households by state food and civil supplies departments and doubles as identity proof. Lacks Aadhaar-level security features, so the same caution applies.

Passport. The most secure document in the mix, with a machine-readable zone, layered print security, and a photograph. New e-passports (2025 rollout) add an embedded chip.

Driver's License. Issued by state transport departments; formats vary across states, but barcodes and QR codes support secure verification. Many states now issue smart-card versions.

Credit / Debit Card — not accepted. Payment cards are not identity documents: no photograph, no biometric anchor, no issuing-authority attestation of who the holder is. Accepting them as identity proof invites fraud — leave them out of your document list.

2.3 | Document Type Preference and Performance

Different Indian documents produce different verification success rates. The grid below pairs each document with its performance profile and the fields Shufti extracts from it.



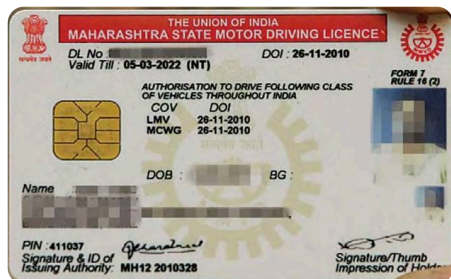
An Aadhaar card specimen with the holder's photograph, bilingual Hindi and English text, and a QR code.

Aadhaar Card

The Aadhaar card's clear multilingual layout suits OCR well, and Shufti's extraction runs faster and cleaner on it than on any other Indian document.

Format: Paper letter, e-document (e-Aadhaar PDF), PVC card

Extraction parameters: Name, date of birth, document number, address, gender, issue date



A Maharashtra State motor driving licence, one of the many state formats in circulation.

Driver's License

Indian driver's licenses are complex, with diverse formats and regional variations, and Shufti's OCR is tuned specifically for these differences.

Format: Physical card, e-document

Extraction parameters: Name, date of birth, issue date, expiry date, address



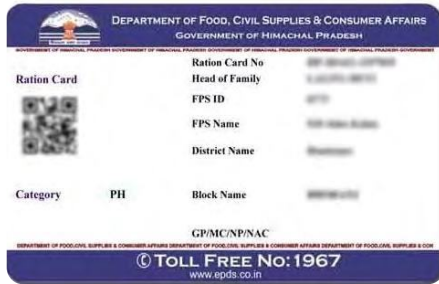
A PAN card specimen from the Income Tax Department with the permanent account number field.

PAN Card

The PAN card is heavily used for financial and administrative checks, and Shufti's systems are optimised for it, returning fast and accurate reads.

Format: Physical card, e-document

Extraction parameters: Name, date of birth, PAN number, issue date, father's name



A Himachal Pradesh ration card with the fair price shop details and a QR code.

Ration Card

Ration cards vary by state and language, and Shufti's OCR processes both physical and electronic versions while adapting to the regional differences.

Format: Physical card, e-document

Extraction parameters: Name, document number, issue date, address, expiry date



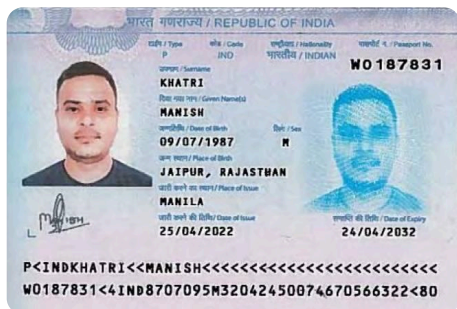
An Election Commission voter ID card, front, with the holder's photograph and elector number.

Voter ID Card

Voter ID cards appear in verification flows across many sectors, and Shufti's OCR handles their regional variations and printed elements accurately.

Format: Physical card, paper-based, e-document

Extraction parameters: Name, date of birth, document number, issue date, address, gender, father's or husband's name



An Indian passport data page with the photograph, bilingual field labels, and the machine readable zone.

Passport

The Indian passport's holograms and microtext challenge ordinary OCR, and Shufti's systems are tuned to capture every field with precision, including the MRZ.

Format: Physical booklet, chip-enabled e-passport on new issues

Extraction parameters: Name, date of birth, issue date, document number, expiry date, nationality, gender, place of issue, MRZ

2.4 | Preferred Documents for Identity Verification

Shufti recognises the Aadhaar card and the PAN card as the preferred documents for identity verification in India. The Aadhaar card leads on the strength of its wide use across sectors and its link to biometric data, while the PAN card is the natural choice for financial and tax-related checks. Shufti's OCR is optimised for the specific features of both, so extraction stays quick and accurate.

Verification Settings

Five settings that decide how well your India verification flow performs. Once you know which documents to accept, these settings shape how the checks behave in production. Set them once for the Indian market and revisit them when formats change.

01

Verification Mode (Onsite vs. Offsite)

Onsite verification is the right default for high-risk industries and compliance-heavy scenarios in India: it runs in real time, confirms each customer with their consent, and adds biometric checks to the document read. Offsite verification suits telecom, e-commerce, and micro-finance, where documents such as PAN cards, voter IDs, or driver's licenses are collected and verified in batches,

02

Data Extraction

Accurate OCR decides whether Indian documents verify at all, given scripts such as Devanagari, Tamil, and Bengali across the same document classes. Shufti's OCR holds its accuracy across these languages and across the document formats each state produces.

03

Identity Document Backside Proof

Indian documents carry material information on both sides. The back of an Aadhaar card, a driver's license, or a voter card can hold the address, the document number, and a barcode or QR code. Require the backside during capture and your extraction stays complete.

04

Document Format Settings for the Aadhaar Card

The Aadhaar card reaches you in two main forms, a laminated paper original and the e-Aadhaar, an electronically generated PDF. Each needs its own handling, since lamination adds glare on camera capture while the e-Aadhaar arrives as a clean digital file. Allow e-documents in your verification settings and customers can submit the e-Aadhaar directly,

5 | Real-Time Proof Capture vs. Proof Upload

Indian identity documents can be verified through two proof methods, and each suits a different situation.

Real-Time Proof Capture. Shufti's auto-capture detects the document in the camera frame and captures it without user effort, the fastest route for physical documents such as the paper Aadhaar card or a voter card.

Proof Upload. Lets users submit digital copies they already hold, such as an e-Aadhaar or a digital PAN card, through a simple upload step.

Merchants control both routes through **Allow Online** (real-time proof capture) and **Allow Offline** (proof upload) in the Verification Settings tab. Offering both keeps completion rates high across India's mix of paper documents and phone-based credentials.

Disclaimer: While we endeavour to keep the information up-to-date and/or correct, we make no representations or warranties of any kind, express or implied, or for any purpose about the completeness, accuracy, reliability, suitability, or availability of the contents or information herein. Any reliance on its content is thus entirely at your own risk. Shufti will not be liable for any false, inaccurate, inappropriate, or incomplete information presented herein, and all liabilities with respect to actions taken, or not taken, based on the contents or information herein, or for any loss sustained by you as a consequence are hereby expressly disclaimed by us.



Global reach, local precision

Shufti verifies identities across 240+ countries and territories and reads each market's documents with local accuracy. In **India** that means the old paper RG, the blue CPF card, every CNH format, and the new national identity card are all handled correctly, so you onboard real people fast and keep fraud out.

[Book a Customised Demo](#)

[See Supported Countries](#)

Shufti is based in the United Kingdom and operates offices worldwide.



United States

16200 SW Pacific
Hwy, Suite H PMB
1053,
Tigard, OR



Cyprus

Arch. Makarios III
Avenue 229,
Limassol



Dubai

Unit 507, Level 5,
Gate District
Precinct
Building 03, DIFC,
Dubai



Hong Kong

8 Queen's Road
East, Wan Chai,
Hong Kong



Pakistan

322-G3 Main Blvd,
Johar Town, Lahore,
Pakistan



Singapore

68 Circular Road
#02-01, Singapore



United Kingdom

Office 408,
Coppergate
House, 10 Whites
Row, London E1
7NF

shuftipro.com

sales@shuftipro.com